DOCUMENT RESUME

ED 155 217 TM 007 346

AUTHOR Marsh, Herbert W.
TITLE Students' Evaluations of Instructional Effectiveness: Relationship to Student, Course and Instructor Characteristics.
PUB DATE Mar 78
NOTE 45p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Canada, March 27-31, 1978) ; Some parts of appendixes may be marginally legible

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Bias; *College Teachers; Correlation; Course Evaluation; Evaluation; Evaluation Criteria; Higher Education; Multiple Regression Analysis; *Predictor Variables; Program Attitudes; Questionnaires; Statistical Analysis; *Student Characteristics; *Student Evaluation of Teacher Performance; Teacher Evaluation; *Test Validity; *Undergraduate Students

ABSTRACT

To determine the relationship between 16 background variables and students' evaluations of instruction, a questionnaire was completed in 511 undergraduate courses at the University of Southern California, Los Angeles. Student variables, including grade point average, class size, expected grade, and prior subject interest, rarely explained 10% of the variance in any student ratings and generally explained less than 5%. Different statistical techniques, however, suggested that 12% to 14% of the variance in the student ratings could be predicted by the set of background variables. The variables most important in predicting evaluations were prior subject interest, expected grade, workload/difficulty, and perhaps, percent taking course for interest only. Of these, prior subject interest was the most important and was better interpreted as a variable affecting quality of education. Background variables did have a small relationship to the evaluations, but results argue against bias interpretation. Workload/difficulty was correlated in the opposite direction as would be expected from a bias effect. No single variable was related to a majority of the evaluation scores. The scores most likely to be biased (overall rating and instructor enthusiasm) were not the scores most related to the background variables. A sample summary of the instructor's rating, and the questionnaire are appended. (Author/JAG)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED155217

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Students' Evaluations of Instructional Effectiveness: Relationship to Student, Course, and Instructor Characteristics

This paper was the basis of a paper
presented at the
Annual Meeting of the American Educational Research Association
Toronto, March, 1978.

Herbert W. Marsh

University of Southern California

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Herbert W. Marsh

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

TM007 346

Running Head: Background Characteristics

ABSTRACT

The purpose of this study was to determine the relationship between 16 background variables (GPA, Class Size, Expected Grade, Prior Subject Interest, etc.) and students' evaluations of instruction in 511 undergraduate courses. Individual background variables rarely explained even 10% of the variance in any student ratings and generally explained less than 5%. Little or no non-linearity was found. Different statistical techniques, multiple regression and canonical correlation, suggested that 12% to 14% of the variance in the student ratings could be predicted by the set of background variables. Three, or perhaps four, background variables were most important in predicting students' evaluations: Prior Subject Interest, Expected Grade, Workload/Difficulty, and, perhaps, Percent Taking Course for General Interest Only. Of these, Prior Subject Interest was the most important. While the background variables did have a small relationship to the evaluations, a host of considerations argues against a simple bias interpretation. Workload/Difficulty was correlated in the opposite direction as would be expected with a bias effect. Also, Prior Subject Interest was better interpreted as a variable impacting quality of education. Furthermore, no one background variable was related to even a majority of the evaluation scores, the effect of the background variables varied dramatically for different evaluation scores, and those evaluation scores most likely to be subject to bias (Overall Instructor Rating and Instructor Enthusiasm) were not the ones most related to the background variables. Only Expected Grade could reasonably be considered a bias, and even this interpretation was subject to alternative explanations.

In spite of the widespread use of students' evaluations as one measure of effective teaching, there is often the fear or suspicion that the ratings lack validity and are adversely affected by variables unrelated to the quality of instruction. The harshest critics even suggest that an instructor need only give high grades and demand little work of students in order to receive high evaluations. The purpose of this paper is to investigate the relationship between different dimensions of students' evaluations and a set of background variables characterising the student, the course, and the instructor.

McKeachie (1973), after reviewing a broad spectrum of student evaluation literature, concluded that a number of potential sources of bias apparently are of little consequence. Remmers (1963), describing a quarter of a century of research with the Purdue Rating Scale for Instructors, also concluded that the ratings are little affected by student/course/instructor characteristics. Hildebrand, Wilson and Dienst (1971) found no correlation between students' evaluations and 10 student/course/instructor characteristics. Menges (Menges, 1973; Costin, Greenough and Menges, 1971) also suggested that potential sources of bias have little effect on ratings, but indicated the need for further study of expected grades, prior student interest in the subject, class size, and reason for taking the course. Overall, Marsh and Kesler (1977) reported curvilinear relationships between class size and several evaluation dimensions, but found the effect to be large for only a Group Interaction dimension. Marsh, Overall and Thomas (1976) found that both expected grade and prior interest in the subject correlated with students' evaluations, but other background variables generally did not. In extensive literature reviews on the effect of expected grades (Marsh, Overall and Thomas, 1976; Feldman, 1976), it was concluded that there is generally a small correlation between expected grades and students'

evaluations, but that a bias in the ratings was only one possible explanation. Feldman (1976) also reported that interest in the subject was correlated with ratings and may explain some of the relationship between expected grades and evaluations; more interest in the subject leads to higher grades and better teaching. In summary, no one background variable was found to have a consistently strong relationship with students' evaluations, but several--particularly Expected Grades, Prior Subject Interest, Class Size and, perhaps, Reason for Taking the Course--were found to have small to moderate correlations in a number of different studies.

Fewer studies have looked at the combined effect of an entire set of background variables on one or more student evaluation items. An important problem in this approach is a careful determination of what are appropriate background variables and what are really evaluation items. For example, Price and Magoon (1971) reported that a set of 11 background variables explained over 20% of the variance in 24 evaluation items. However, student ratings of "availability of the instructor," "explicitness of course policies," and "classroom atmosphere" (relaxed versus tense) were considered as background variables and contributed to the prediction of the evaluation items. Most researchers would consider these variables to be part of the evaluation of teaching. Similarly, Pohlmann (1975) found that 9 background variables explained over 20% of the variance in 5 rating items; however, course difficulty was the rating item best predicted, and it was correlated to a conceptually-similar item concerning the hours spent outside of class. Brown (1974) reported that 11 background variables explained 14% of the variance in an average of student evaluation items, but indicated that average grade accounted for the most variance. Burton (1975) showed that 8 background variables (including GPA, Expected Grade, Class Standing, Reason for Enrolling, and Enthusiasm toward

the Subject) explained between 8% and 15% of the variance in instructor ratings over a seven-semester period of time, but indicated that the most important variable was student enthusiasm toward the subject.

In summary, studies considering the combined effect of an entire set of background variables generally found that the set explained at least 10% of the variance in students' evaluations, and some suggest that the proportion was as high as 25%. However, particularly with those studies finding the higher estimates, there was a problem in determining what were background variables and what were evaluation items: either evaluation-like items were included in the set of background items, or conceptually-similar items (e.g., Hours Required and Course Difficulty) were included in both sets.

A host of philosophical and methodological considerations complicates the analysis of the relationship between students' evaluations and student/course/instructor variables. First, correlations cannot be used to prove causation. If poor teachers were assigned to teach large introductory courses, their lower ratings should not be attributed to a bias produced by class size or course level. Second, the distinction between practical and statistical significance needs to be drawn. A statistically significant relationship based upon a very large sample size may be so small as to be of no practical importance. Third, the existence of curvilinear relationships needs to be explored. Fourth, the multivariate nature of students' evaluations requires that different evaluation dimensions be considered separately; class size is moderately correlated with Group Interaction but shows little relationship to other dimensions. Fifth, the combined effect of an entire set of background variables needs to be determined as well as the effect of each separately. Finally, the nature of the "bias" being considered needs careful attention. On the one hand, if

teachers need only give high grades and demand little work to receive high evaluations, then the evaluations are clearly biased. On the other hand, if students start a class with a strong interest in the subject, they may rate the teacher more favorably because he really was more effective than he would have been with less motivated students. In summary, the complications make the problem interesting, but virtually eliminate the possibility of reaching any definitive conclusions.

METHOD

Evaluation Instrument

The evaluation instrument (see Appendix I) consisted of 35 evaluation items which define 9 different evaluation factors, and 6 additional items which measure background variables. Both the individual evaluation items and the evaluation factors are quite reliable (see Appendix II). Coefficient alphas (Nie, et al., 1977) for the evaluation factors varied between .88 and .97. Factor analysis (see Appendix III) has supported the existence of the nine evaluation factors in each of three different semesters. The oblique factor solution resulted in factors which had low to moderate correlations varying between r=-.02 to r=+.49 (median r=.27). Furthermore, essentially the same evaluation factors were found with faculty self-evaluations of their own teaching when using the same instrument (Marsh and Overall, 1978).

Students' evaluations were summarized by eleven evaluation scores, the nine evaluation factors and the two overall summary items. Factor scores were weighted averages of the evaluation items, while the two overall ratings

were based upon responses to single items. The evaluation scores and brief descriptions are as follows:

Learning/Value--The extent to which students felt they encountered a valuable learning experience that was intellectually challenging.

Instructor Enthusiasm--The exteni to which students perceived the instructor to display enthusiasm, energy, humor and an ability to hold interest.

Organization--The instructor's organization of the course, course materials, and class presentations.

Group Interaction--Students' perceptions of the degree to which the instructor encouraged class discussions and invited students to share their own ideas or be critical of those presented by the instructor.

Individual Rapport--The extent to which students perceived the instructor to be friendly, interested in students, and accessible in or out of class.

Breadth of Coverage--The extent to which students perceived the instructor to present alternative approaches to the subject and to emphasize analytic ability and conceptual understanding.

Examinations--Students' perceptions of the value and fairness of graded materials in the course.

Assignments--The value of class assignments (readings, homework, etc.) in adding appreciation and understanding to the subject.

Workload/Difficulty--Students' perceptions of the relative difficulty/workload of the course and the pace of presentations.

Overall Course--A single item asking students to compare the course with other courses at USC.

Overall Instructor--A single item asking students to compare the instructor with other instructors at USC.

Background Variables

The set of background variables consisted of 16 different variables describing the course, students in the course, and the instructor. Selection of this set of variables was prompted by a review of the literature and the availability of information. The Workload/Difficulty variable has been included as both an evaluation score and a background variable. However, whenever the combined effect of background variables was being determined, Workload/Difficulty was considered a background variable.

Two subsets of the background variables were given special attention in some of the analyses. The first subset was the Reason for Taking the Course. Students selected one of five possible reasons for taking the course, or left the item blank if none of the given reasons were appropriate. The percentages of students indicating each of the reasons were included as five separate variables. However, these five variables should be interpreted cautiously, since a higher percentage for any one necessitates a lower percentage on another. The second subset consisted of four highly correlated variables which describe the Class Level.

The 16 background variables and brief descriptions are as follows:

Level of Interest in the Subject Prior to This Course (1-Very Low... 3-Medium...5-Very High)--Mean class average response was 3.4.

Workload/Difficulty--An evaluation factor score representing four items; high values refer to courses which are more difficult, have a heavier workload, are faster paced, and require a greater number of hours outside of class.

Overall GPA (1-Below 2.5, 2-2.5 to 3.0, 3-3.0 to 3.4, 4-3.4 to 3.7, 5-Above 3.7)--Mean class average response was 3.3 (i.e., slightly higher than a B average.

Enrollment--The number of students who were enrolled in the course (mean enrollment was 34.5).

Teacher Rank (1-Teaching Assistant, 2-Lecturer, 3-Instructor, 4-Assistant Professor, 5-Associate Professor, 6-Full Professor)--Class average response was 4.3 (Note: Teaching Assistants were excluded from the analysis, and very few teachers were either Lecturers or Instructors).

Percent Students Majoring in Same Division as the Course (e.g., % Social Science students in Social Science courses)--Mean class average response was 49%.

Expected Grade (1-F, 2-D, 3-C, 4-B, 5-A)--Mean class average response was 4.2 (i.e., slightly higher than a B average).

[1]Reason for Taking the Course (a subset of five variables):

% indicating Major Requirement--Mean class average response was 42%.

% indicating Major Elective--Mean class average response was 24%.

% indicating General Interest Only--Mean class average response was 16%.

% indicating General Education Requirement--Mean class average response was 12%.

% indicating Minor/Related Field--Mean class average response was 5%.

Class Level (a subset of four variables):

Mean Year in School (1-Freshman, 2-Sophomore, 3-Junior, 4-Senior, 5-Graduate)--Mean class average response was 3.2.

% indicating Freshman or Sophomore--Mean class average was 25%.

% indicating Junior or Senior--Mean class average response was 65%.

Course Level (1-Lower Division, 2-Upper Division)--Mean class average response was 1.7.

Statistical Analysis

All analyses were performed on class average responses for the 511 courses in the study. Each of the 16 background variables was correlated with each of the 11 student evaluation scores. These linear relationships were considered substantial only if they predicted at least 5% of the variance in one of the evaluation scores (i.e., the correlation was at least r=+.23). Second order (quadratic) and third order (cubic) components of each background variable

---

[1]Students were also given the option of leaving this item blank; thus, there is an implicit sixth possible response category of "Other" which keeps this set from being completely dependent.

were then tested to determine if any substantial non-linearity existed, a relationship being considered substantially non-linear if it accounted for at least 5% of the variance in an evaluation score and if a non-linear component added at least 1% to the variance explained by the linear relationship.

Step-wise multiple regression (Nie, et al., 1975) was used to determine the combined effect of the background variables on each evaluation score. At each step, the single variable which added most to the "variance explained" was added to the regression equation until no additional variable could add an additional 1% to the total variance already explained. At each step, the total variance explained was adjusted for the number of variables in the equation (Nie, et al., 1975; Cohen and Cohen, 1975). Cohen and Cohen (1975) suggested that all variance estimates be corrected for the total number of variables which are available to be used in the regression equation rather than just the number used at each step; this adjustment procedure was the basis of final variance estimates.

The proportion of variance which was uniquely contributed by each of the background variables was then obtained by determining the proportion of variance which could be predicted by all but one of the background variables and then computing the proportion of additional variance which could be explained by the one remaining variable. A variation of this procedure was used for the two subsets of background variables discussed earlier (Reason for Taking Course and Class Level). The multi-colinearity (Cohen and Cohen, 1975) dictated that any one variable in each set would make little contribution once the remaining variables had been included. To avoid this problem, two alternatives were considered. First, the additional variance contributed by the entire subset of

variables was determined. Second, the additional variance contributed by each variable in the subset was determined without including any of the other variables in the subset. Once again, the proportion of variance explained was adjusted for the number of variables included at each step.

Canonical correlation (Dixon, 1975; Cooley and Lohnes, 1971, 1976) was used to determine the relationship between the entire set of background variables and the entire set of student evaluation scores. In the first step, this procedure determines a linear combination of background variables which is maximally correlated with a linear combination of student evaluation scores. At each successive step, additional linear combinations of variables, the canonical variates, are extracted which are uncorrelated with previous ones and maximally correlated with each other. This procedure is intuitively an extension of multiple regression in which a linear combination of background variables was determined which maximally correlated with just one evaluation score. Cooley and Lohnes (1971, 1976) have discussed a measure of the redundancy between two sets of variables used in canonical correlation. This measure is a quantitative description of the total proportion of variance in one set of variables which can be predicted by another. In this study, the redundancy measure was used to determine the proportion of variance in the entire set of evaluation scores which can be predicted by the entire set of background variables.

RESULTS AND DISCUSSION

Bivariate Relationships--Linear and Non-Linear

Linear correlations between each of the 16 background variables and the 11 evaluation scores are presented in Table 1. Correlations as small as r=.09 were statistically significant, though of little practical significance. Consequently, attention was focused upon those relationships which accounted for at least 5% of the variance in any one of the evaluation scores (i.e., correlations of at least r=.23). Of the 175 correlations, only 18 met this criteria, and only 3 of the 18 accounted for as much as 10% of the variance: Prior Subject Interest was positively correlated with both Overall Course Rating and Learning/Value evaluation scores, and Enrollment was negatively correlated with quality of Group Interaction. None of the 16 background variables accounted for more than 5% of the variance in even a majority of the evaluation scores, and only three (Prior Subject Interest, Workload/Difficulty, and Expected Grade) did so for more than one evaluation score.

INSERT TABLE 1 ABOUT HERE

The extent of non-linearity was considered in each of the background-evaluation relationships. Only 7 of 175 relationships showed any substantial non-linearity. Course enrollment generally showed a non-linear relationship to evaluations: courses with large enrollments and small enrollments were rated more favorably, but only 2 of the 11 relationships reached the criteria of

"substantial." For the other nine evaluation scores, even with the additional variance accounted for by the non-linear components of enrollment, the total variance explained was less than 5%. The majority of the evaluation scores (Overall Course, Overall Instructor, Enthusiasm, Organization, Individual Rapport, and Examinations) showed no substantial non-linear relationship with any of the background variables.

In summary:

**Individual background variables rarely accounted for as much as 10% of the variance in any of the evaluation scores (3 of 175 relationships) and generally did not even account for 5% (18 of 175).

**More favorable evaluations tended to be given to classes in which students had higher Prior Subject Interest and Expected Grades and those in which they experienced higher levels of Workload/Difficulty.

**Background variables generally showed little or no non-linear relationship to any of the evaluation scores.

Multivariate Relationships--Multiple Regression

Each of the set of 16 background variables was entered into a step-wise multiple regression to predict each of the 11 student evaluation scores. This analysis had two purposes: to determine the combined effect of all the background variables on each evaluation score, and to determine which of the background variables consistently made the largest contribution. In order to simplify the interpretations, a rather conservative criterion was used to determine whether additional variables would be entered into the equation; an

additional variable was included only if it added at least 1% to the variance which had been accounted for already by the previous set of variables.

The percentage of variance which could be explained in each of the different evaluation scores (corrected for the number of background variables available) varied dramatically, ranging from 0% for Organization to 25% for Learning/Value. The set of background variables accounted for 20% or more of the variance in three evaluation scores: Overall Course (20%), Group Interaction (23%), and Learning/Value (25%). Four of the set of background variables consistently appeared in the final regression equations: Prior Subject Interest, Expected Course Grade, Workload/Difficulty, and percentage of students indicating "General Interest" as their reason for enrolling in the course (as opposed to Major Requirement, Major Elective, or General Education Requirement). In each case, courses tended to be rated more favorably when Prior Subject Interest was higher, when Workload/Difficulty was greater, when Expected Grades were higher, and when percent enrolling for "General Interest" was higher.

INSERT TABLE 2 ABOUT HERE

Multiple regression was also used to determine the unique contribution of each of the individual background variables to each of the evaluation scores (see Table 3). The unique variance is the proportion of additional variance accounted for by each variable or each subset of variables, after all other variables have been considered. Inspection of Table 3 indicates that much of the variance accounted for by any one background variable is redundant with variance accounted for by others; although 18 of 175 relationships between

background variables and evaluation scores accounted for 5% of the variance, only 7 relationships uniquely accounted for at least 5% (i.e., variance which was not also explained by other background variables). Only four Background variables uniquely accounted for as much as 5% of the variance in any of the evaluation scores: Prior Subject Interest, Workload/Difficulty, Expected Grade, and Reason For Taking Course.

INSERT TABLE 3 ABOUT HERE

An interesting relationship consistently appeared between Expected Grade, Workload/Difficulty, and the evaluation scores. Expected Grade and Workload/Difficulty both tended to be positively related to each of the evaluation scores, but were negatively related to each other (r=-.29). This rather unusual event is a case of cooperative or reciprocal suppression which is described by Cohen and Cohen (1975). While this occurrence is interesting, it also complicates interpretations. The combined effect of the two variables is necessarily greater than the sum of their individual effects. Furthermore, the supposedly "unique" variation attributable to either variable may be greater than variance explained before the effect of other variables has been removed.

In summary:

**Percentage of variance in different evaluation scores which was explained by the set of background variables varied dramatically, ranging from 0% to 25%; the average was 11.8%.

**Background variables explained 20% or more of the variance in 3 evaluation scores: Learning/Value, Group Interaction, and Overall Course.

**Only four of the background variables consistently appeared in the final regression equations: Prior Subject Interest, Expected Grade, Workload/Difficulty, and percentage indicating "General Interest Only" as the reason for taking the course.

**Much of the variance predicted by individual background variables was redundant with variance predicted by other background variables.

Multivariate Relationships--Canonical Correlation

Canonical correlation is a general statistical technique for determining the combined effect of one set of variables (the background variables) on another set of variables (the evaluation scores). When only one evaluation score is considered, the technique is equivalent to multiple regression. When more than one evaluation score is considered, linear combinations of the background variables and the evaluation scores are determined so that their correlation is maximal. On successive steps in the analysis, additional pairs of linear combinations are extracted which are again maximally correlated with each other and uncorrelated with previous linear combinations. The magnitude of the canonical correlations can easily be misinterpreted. For example, if the same variable were included in both sets, the first canonical correlation would necessarily by r=1.0--a perfect relationship. Each canonical variate would consist of only the one variable in common to the two sets, and all other variables would have a zero weighting. Cooley and Lohnes (1971, 1976) describe a redundancy measure which indicates how much variance in one set of variables can be predicted by a second set of variables. If Workload/Difficulty were included as both a background variable and an evaluation score, the first

canonical correlation would be r=1.0, but the redundancy measure would be only 9% (e.g., 1 of the 11 evaluation scores would be perfectly predicted, but the other 10 would not be predicted at all).

Inspection of Table 4 indicates that 9.5% of the variance in the evaluation scores can be explained by the first pair of canonical variates. Successive canonical variates account for 2.3%, 0.8%, 3.2% and 0.5% of the variance in the evaluation scores. The sum of these values, approximately 16%, is an estimate of the variance in the entire set of students' evaluations which can be explained by the entire set of background variables. This sample estimate is inflated in the same way as the multiple R^2 in multiple regression, but no adjustment procedure to correct for this positive bias has been developed. If this variance estimate were corrected for the use of 16 background variables in the same manner as the multiple R^2 (Cohen and Cohen, 1975), the corrected estimate would be 13.6%. This result, of course, corresponds rather closely to the 11.8% value which was based upon the successive multiple regression equations. Furthermore, Prior Subject Interest, Workload/Difficulty, and Expected Grade were again found to be the most important in predicting the evaluation scores; while Learning/Value, Group interaction, and Overall Course Rating were the evaluation scores which were best predicted. However, Teacher Rank was also shown to be important in this particular analysis; higher ranked teachers (e.g., Full Professors) were rated as giving broader coverage, better assignments, and poorer Group Interaction.

INSERT TABLE 4 ABOUT HERE

In summary:

**Canonical correlation indicated that 16% of the variance in the entire set of evaluation scores can be explained by the entire set of background variables.

**Background variables most important in explaining evaluation scores were Prior Subject Interest, Workload/Difficulty, Expected Grade, and Teacher Rank.

**Evaluation scores which were best predicted were Learning/Value, Group Interaction, and Overall Course Rating.

Most Important Background Variables and Explanatory Models

A difficult problem in multivariate research is the search for "the" most important independent variables (the background variables) used in predicting one or more dependent variables (the evaluation scores). Many alternative criteria, including their strengths and weaknesses, have been discussed elsewhere, but the most commonly suggested are: 1) simple correlations, 2) Beta weights in final regression equations, 3) the change in variance explained as each new variable enters a regression equation, and 4) the variance which is uniquely defined by each independent variable. Fortunately, in the present situation, each of these criteria suggested that the same three, or perhaps four, background variables were most important: Prior Subject Interest, Workload/Difficulty, Expected Grade, and perhaps Reason for Taking the Course. Most favorably rated courses tended to have students who were more interested in the subject before the start of the course, tended to have students who

expected to receive higher grades, tended to have a heavier workload and be more difficult, and tended to have more students taking the course for general interest only.

Path analysis, commonality modeling, and related techniques of variance partitioning (Cooley and Lohnes, 1976; Blalock, 1971) attempt to determine the relative contribution of each independent variable and how it affects the dependent variables. While none of these techniques allow the researcher to draw causal conclusions on the basis of correlational data, they do provide a systematic approach to testing some causal hypotheses. Generally, the first step is to establish the temporal ordering of the independent (background) variables, and then to determine the proportion of variance directly attributable to each. For example, Feldman (1976) speculated that the observed relationship between grades and evaluations may be partially or fully due to the fact that both are causally related to some antecedent variable such as prior interest. He proposed a path analysis which would test this hypothesis, but indicated that none of the studies which he reviewed provided any test of this hypothesis.

In this study, a temporal ordering of the four most important background variables was established. Prior Subject Interest and Reason for Taking the Course (represented by percent indicating "General Interest Only") were assumed to come first. Since these two variables were essentially uncorrelated ($r=+.05$, not statistically significant), their combined effect was approximately the sum of the separate effects of each. Expected Grade and Workload/Difficulty were assumed to come later, and so only variance not already explained by the first two variables was attributed to them. However, as discussed earlier,

Expected Grade and Workload/Difficulty were negatively correlated with each other ($r=-.29$, $p < .001$) even though both were generally positively related to the evaluation scores. This complicates their interpretation since their combined contribution was greater than the sum of each of them separately.

Inspection of Table 5 indicates that, averaged across all evaluation scores, 12.5% of the variance in the evaluation scores can be explained by the four background variables. This value, if corrected for the four background variables included in each regression, would be 11.9%. Comparison with Table 2 indicates that only Group Interaction (which was related to Enrollment) and Breadth of Coverage (which was related to Teacher Rank and percent taking the course as a Major Requirement) were substantially better predicted by the entire set of 16 background variables. Almost half of the variance which was explained by the four background variables was directly attributable to Prior Subject Interest (5.3% of 12.5%), while each of the other variables contributed no more than 20% of the predictable variance (i.e., no more than 2.5% of the 12.5%). Controlling for the effect of Prior Subject Interest and Reason for Taking the Course had little effect on the variance explained by Workload/Difficulty, but reduced the effect of Expected Grade by nearly one-third. As previously noted, the combined effect of Expected Grade and Workload/Difficulty (5.7% of 12.5%) was greater than the sum of each separately.

INSERT TABLE 4 ABOUT HERE

In summary:

**A variety of criteria all indicated that Prior Subject Interest, Workload/Difficulty, Expected Grade, and Percent Taking Course for General Interest Only were the most important background variables in predicting the evaluation scores.

**Averaged across all the evaluation scores, these four background variables explained 12.5% of the variance (11.9% when corrected).

**The explanatory model used in this analysis indicated that Prior Subject Interest was the most important background variable, accounting for almost half the predictable variance.

**Controlling for the effect of Prior Subject Interest reduced the effect of Expected Grade by one-third, but had little effect on the other background variables.

CONCLUSIONS AND IMPLICATIONS

The purpose of the study was to describe the relationship between a set of students' evaluations of instructional effectiveness and a set of 16 background variables describing the student, the course, and the instructor. Individual background variables rarely explained as much as 10% of the variance in any of the evaluation scores (3 of 175 relationships) and generally did not even explain 5% (18 of 175). Multiple regression indicated that the percentage of variance explained in different evaluation scores varied dramatically, ranging from 0% to 25%; the average across the 10 evaluation scores was 11.8%.

Canonical correlation showed that 16% of the variance in the students' evaluations could be explained by the set of background variables, but this estimate was known to have a slight positive bias.

A variety of criteria each suggested that three or perhaps four background variables were most important in explaining variance in the evaluation scores: Prior Subject Interest, Workload/Difficulty, Expected Grade, and, perhaps, Reason for Taking the Course. Depending upon what assumptions were made, these four accounted for either all or more than 80% of the variance attributable to the entire set of 16 background variables. An explanatory model based upon these four background variables indicated that almost half the predictable variance in the evaluation scores could be explained by Prior Subject Interest alone. Controlling for the effect of Prior Subject Interest reduced the variance attributable to Expected Grade by one-third, but had little effect on the other two background variables.

In summary, the relationship between the 16 background variables and the set of students' evaluations definitely existed, but tended to be rather small. Different statistical procedures suggested that, on the average, 12% to 14% of the variance in students' evaluations could be explained by the entire set of background variables. Even if this entire relationship were assumed to be due to biases in the students' evaluations, the magnitude of this bias would not obviate their usefulness. Many well-accepted psychological tests probably have biases as large as this or larger, even if the people being evaluated are not as clever at identifying the biases as are university faculty. However, a host of arguments suggest that even the relationship which was found should not be considered a simple bias in the students' evaluations.

At the most simplistic level, critics of students' evaluations have suggested that instructors need only give high grades and demand little work of students in order to be favorably evaluated. If this were true, the bias would affect each of the evaluation scores in a similar manner. This simplistic notion of bias can clearly be rejected. No one background variable was substantially related to even half the evaluation scores, and the percentage of variance in the different evaluation scores which was explained by the set of background variables varied dramatically. Furthermore, the evaluation scores which would be expected to be most subject to bias were not. More than twice as much variance was explained in Overall Course Rating than in Overall Instructor Rating, and four times as much variance was explained in Learning/Value than in Instructor Enthusiasm. If students' evaluations are biased, the bias is not a simple one.

A detailed inspection of the background variables most related to the evaluation scores also undermined the speculation that the relationship is caused by a simple bias. The Workload/Difficulty variable was related to the evaluation scores in the opposite direction from what would be predicted by a bias. Harder, more difficult courses which required more time outside of class were rated more favorably. Prior Subject Interest can be better interpreted as a variable impacting the quality of education than a bias which is specific to students' evaluations. A high Prior Subject Interest creates a more favorable learning environment and probably makes it easier to do a more effective job of teaching. Furthermore, Marsh and Overall (1978) showed that this variable effected faculty self-evaluations of their own teaching as well as the students' evaluations.

The only background variable which could reasonably be considered to be a possible bias to students' evaluations is Expected Grade, and even this interpretation is subject to alternative explanations. First, the effect of Expected Grade was reduced by one-third by controlling for the effect of Prior Subject Interest. The most plausible explanation is that higher Prior Subject Interest causes both better grades and a better educational experience. Second, the Expected Grade relationship can only be considered a bias if the higher grades reflect "easy grading" on the part of the teacher. If higher expected grades reflect actual student achievement--better students' evaluations are related to better student learning--then most researchers would interpret the relationship to support the validity of the ratings. In reality, Expected Grade probably reflects some unknown combination of both "easy grading" and student achievement.

In conclusion, a variety of multivariate techniques suggested that 12% to 14% of the variance could be predicted by a set of 16 background variables. However, even this small to moderate relationship could not be interpreted as a simple bias in the students' evaluations. No one background variable was substantially related to even a majority of the evaluation scores, the percentage of variance explained in different evaluation scores varied dramatically, and the evaluation scores most likely to be subject to bias were not the ones best explained by the background variables. Of the three background variables most clearly related to the students' evaluations, one (Workload/Difficulty) was correlated in the opposite direction from what would be predicted by a bias hypothesis, and another could be better interpreted as affecting quality of teaching rather than a bias. Only Expected Grade could reasonably be considered a bias, and even this interpretation was subject to alternative explanations.

REFERENCES

Blalock, H.M. (Ed.) Causal Models in the Social Sciences. Chicago: Adine Atherton, 1971.

Brown, D.L. Faculty Ratings and Student Grades: A Large-Scale Multivariate Analysis By Course Sections. An unpublished Dissertation, 1974.

Burton, D. Student ratings-an information source for decision-making. In R.G. Cope (Ed.) Information for Decisions in Postsecondary Education: Proceedings of the 15th Annual Forum, Association for Institutional Research, St. Louis, Missouri, April, 1975.

Centra, J.A. The Student Instructional Report: Item Reliabilities (Student Instructional Report No. 3) Princeton, N.J: Educational Testing Service, 1973.

Cohen, J. and Cohen, P. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. New York: Wiley, 1975.

Cooley, W.W. & Lohnes, P.R. Multivariate Data Analysis. New York: Wiley, 1971.

Cooley, W.W. & Lohnes P.R. Evaluation Research in Education. New York: John Wiley & Sons, 1976.

Costin, F., Greenough, W.T., and Menges, R.J. Student ratings of college teaching: Reliability, validity, and usefulness. Review of Educational Research, 1971, 41, 511-535.

Dixon, W.J. (Ed.) Biomedical Computer Programs. University of California Press, 1973.

Feldman, K.A. Grades and college students' evaluations of their courses and teachers. Research in Higher Education, 1976, 4, 69-111.

Hildebrand, M., Wilson, R.C., and Dienst, E.R. Evaluating University Teaching. Berkeley: Center for Research and Development in Higher Education, University of California, Berkeley, 1971.

Marsh, H.W. Reliabilities of Items on the Student Evaluation Form Used in the Social Science Division at USC. Los Angeles: Office of Institutional Studies, University of Southern California, 1976 (OIS 76-8)

Marsh, H.W., Overall, J.U., and Kesler, S.P. The Validity of Students' Evaluations of Instructional Effectiveness: A Comparison of Faculty Self-Evaluations and Evaluations By Their Students. Paper presented at Annual Meeting of the Association of Institutional Research, Houston, May, 1978.

Marsh, H.W., Overall, J.U., and Thomas, C.S. The Relationship Between Students' Evaluation of Instruction and Expected Grade. Paper presented at Annual Meeting of American Educational Research Association, San Francisco, April, 1976. (ERIC ED 126 140)

McKeachie, W.J. Correlates of Student Ratings. In Sockloff, A.L. (Ed.), Proceedings: The First Invitational Conference on Faculty Effectiveness as Evaluated by Students. Philadelphia: Measurement and Research Center, Temple University, 1973.

Menges, R.J. The new reporters: students rate instruction. In C.R. Pace (Ed.) Evaluating Learning & Teaching. San Francisco: Jossey-Bass, Inc., 1973.

Nie, N.H., Hull, C.H., Jenkins J.G., Steinbrenner, K. and Bent, D.H. Statistical Package For The Social Sciences. New York: McGraw-Hill Book Company, 1975.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., and Bent, D.H. Update to Statistical Package For the Social Sciences. New York: McGraw-Hill Book Company, 1977.

Overall, J.V., Marsh, H.W. and Kesler, S.P. Class Size and Students' Ratings of Instruction: A Clarification of Relationship. Paper presented at Annual Meeting of American Educational Research Association, New York, April, 1977.

Pohlmann A multivariate analysis of selected class characteristics and student ratings of instruction. Multivariate Behavorial Research, 1975, 10, 81-92.

Price, J.A. and Magoon, A.J. Predictors of College Student Ratings of Instructors. Proceedings of the 79th Annual Convention of the American Psychological Association, 1971, 6, 523-524.

Remmers, H.H. Teaching methods in research on teaching. In N.L. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

Weiner, B.J. Statisical Principles in Experimental Design. New York: McGraw-Hill, 1971.

TABLE ONE

Correlations Between Background/Demographic Variables and Student Evaluation Scores
(N=511 Class Averages)

	Overall Course	Overall Instr.	Learning	Enthu-siam	Organ-ization	Group Inter.	Indiv. Rapport	Breadth	Exams	Assign.	Work/ Diff.
Prior Interest in Subject (1-Low...5-High)	[.33]	.20	[.44]	[.23]	-.03	[.29]	.09	-.03	.03	.20	(.12)
Workload/Difficulty (1-Light/Easy...5-Heavy/Hard)	[+.23]	+.14	+.12	+.06	.01	-.02	.01	.15	.10	[.23][3]	(1.0)
Expected Grade (0-F...4-A)	.21	.20	[.29]	.20	.01	[.31]	.17	-.02	.18	.13	[(-.29)]
Course Level (1-Lower Division, 2-Upper Division)	.17	.14	.21	.12	-.08	[.29]	.14	.13	.04	.11	(.06)
[1]% Taking as"Major Elective"	.16	.13	[.26]	.06	-.03	.21	.04	.18[3]	.02	.15	(-.06)
[1]% Taking as "General Interest Only"	.16	.12	.15	.09	.16	.07	-.02	.19[3]	.10	.18	(-.13)
Overall GPA	.07	.07	.10	.07	-.06	.17	.14	.04	.07	.13	(.12)
[1]% Taking as "Major Requirement"	-.15	-.12	-.18	-.07	-.08	-.04	.01	[-.26][3]	-.02	-.17	(.17)
% Majoring in Division (Social Science)	.15	.14	.15	.03	.05	[.29]	.08	.08	.13	.11	(.15)
% Freshman-Sophomore Students	-.12	-.12	-.18	-.08	-.01	[-.28]	-.17	.01	-.09	-.05	(-.03)
[1]% Taking as a "General Education Requirement"	-.11	-.08	-.17	-.04	.03	[-.28]	-.06	.03	-.09	-.06	(-.12)
Average "Year in School" (1-Freshman...5-Grad.)	.11	.10	.20	.11	-.06	[.27]	.19	-.06	.03	.04	(.03)
Enrollment	-.10	-.09	-.14[3]	.01	-.03	[-.32][3]	-.18	.01	-.13	-.04	(.01)
% Junior-Senior Students	.11	.01	.07	.21	.04	.05[3]	.13	.03	.10	.09	(-.01)
Teacher Rank (1-Lecturer...4-Full Prof.)	-.02	-.08	-.10	-.12	-.10	-.14	-.05	[+.24][3]	-.14	+.13	(.11)
[1]% Taking Course in"Minor/Related Field"	.07	.06	.07	.03	.02	.02	.04	.12	.01	.07	(-.01)
[4] Multiple R^2 (% variance explained)	20.4	8.9	24.7	5.9	0	23.0	3.5	11.3	8.3	12.3	19.6

1- Students each indicated one of these reasons for taking the course
2- Correlations in boxes represent relationships in which a single Background/Demographic Variable accounts for at least 5% of the variance in a student evaluation score
3- These relationships show substantial non-linear trends(i.e. quadratic and/or cubic components add at least 1% to the Variance Explained by the linear relationship and Total Variance Explained was at least 5%)
4- See Table Two for details on how this value was obtained

TABLE TWO

Multivariate Relationship Between Each Student Evaluation Score and the Entire Set of Background/Demographic Variables[4]

(N=511 Class Averages)

OVERALL COURSE RATING

Step #	Variable	ΔR^2	Beta	r
1	Prior Interest	10.4	.23	.32
2	Work/Diff.	3.5	.34	.23
3	Expected Grade	4.9	.32	.21
4	% General Interest	3.4	.18	.16
	Total Variance Accounted	22.2% (20.4%)		

OVERALL INSTRUCTOR RATING

Step #	Variable	ΔR^2	Beta	r
1	Prior Interest	3.9	.12	.20
2	Expected Grade	2.6	.24	.20
3	Work/Diff.	2.9	.21	.14
4	% General Interest	1.6	.13	.12
	Total Variance Accounted	11.1% (..9%)		

LEARNING

Step #	Variable	ΔR^2	Beta	r
1	Prior Interest	.194	.36	.44
2	Expected Grade	.038	.26	.29
3	Work/Diff.	.018	.17	.12
4	% General Interest	.021	.15	.15
	Total Variance Accounted	27.1% (24.7%)		

ORGANIZATION

Step #	Variable	ΔR^2	Beta	r
1	% General Interest	.022	.16	.16
	Total Variance Accounted	2.2% (0%)		

GROUP INTERACTION

Step #	Variable	ΔR^2	Beta	r
1	Enrollment	.102	-.21	-.32
2	Expected Grade	.083	.24	.31
3	Prior Interest	.025	.16	.29
4	% Majoring in Div.	.018	.18	.29
5	% General Interest	.011	.11	.07
	Total Variance Accounted	24.7% (23.0%)		

INDIVIDUAL RAPPORT

Step #	Variable	ΔR^2	Beta	r
1	Avg. Year in School	.036	.11	.19
2	Enrollment	.013	-.12	-.18
3	Expected Grade	.011	.12	.17
	Total Variance Accounted	6.0% (3.5%)		

BREADTH OF COVERAGE

Step #	Variable	ΔR^2	Beta	r
1	% Major Requirement	.068	-.27	-.26
2	Work/Diff.	.039	.18	.15
3	Teacher Rank	.029	.18	.24
	Total Variance Accounted	13.6% (11.3%)		

ENTHUSIASM

Step #	Variance	ΔR^2	Beta	r
1	Prior Interest	.050	.20	.23
2	Expected Grade	.034	.16	.20
	Total Variance Accounted	8.4% (5.9%)		

EXAMINATIONS

Step #	Variable	ΔR^2	Beta	r
1	Expected Grade	.031	.22	.18
2	Work/Diff.	.024	.20	.10
3	Teacher Rank	.018	-.15	-.14
4	PJRSR	.014	.13	.13
5	% General Int.	.016	.13	.10
	Total Variance Accounted	10.3%(8.3%)		

ASSIGNMENTS

Step #	Variable	ΔR^2	Beta	r
1	Work/Diff.	.051	.33	.23
2	% Maj. Require.	.045	-.13	-.17
3	Expected Grade	.038	.21	.13
4	% General Int.	.010	.14	.18
	Total Variance Accounted	14.4% (12.3%)		

WORKLOAD/DIFFICULTY

Step #	Variable	ΔR^2	Beta	r
1	Expected Grade	.087	-.42	-.29
2	Overall GPA	.071	.25	.12
3	Prior Interest	.024	.20	.12
4	% Maj. Require.	.019	.17	.17
5	Teacher Rank	.011	.12	.11
	Total Variance Accounted	21.2% (19.6%)		

1 - Each addtional step was included only if the ΔR^2 (the change in Total Variance Accounted resulting from the Step) was greater than .01

2 - ΔR^2 and Total Variance Accounted have been adjusted for the number of variables included at each step (Nie, et.al.197)

3 - The value of "Total Variance Explained" which appears in parentheses has been corrected for the total number of variables in the set of Background/Demographic variables as is appropriate when using a stepwise regression procedure (Cohen & Cohen, 1975).

4 - The average of the total variance estimates in parentheses is 11.8% (the average of values not in parentheses is 14%) and constitutes one estimate of the proportion of variance in all student evaluation scores which are explained by background variables; Workload/Difficulty was excluded from consideration as an evaluation score in determining the averages.

TABLE THREE

The "Unique" Contribution of Each Background Variable
(N=511 Class Averages)

BACKGROUND/DEMOGRAPHIC VARIABLES	STUDENT EVALUATION SCORES										
	Overall Course	Overall Instructor	Learn	Ethusiasm	Organization	Group Instruction	Individual Rapport	Breadth	Exams	Assignments	Workload/ Difficulty
Prior Subject Interest	2.7%	0.6%	6.7%	1.9%	0.1%	0.9%	0.0%	0.8%	0.3%	3.4%	2.8%
Workload/Difficulty	8.9%	3.8%	3.6%	1.3%	0.2%	0.0%	0.1%	2.1%	2.5%	5.8%	
Expected Grade	5.6%	4.4%	5.0%	3.1%	0.1%	3.8%	2.0%	0.0%	4.5%	1.9%	13.1%
Reason for Taking Course[2]	3.3%	1.8%	3.2%	0.2%	3.4%	1.1%	0.6%	5.9%	1.2%	3.9%	2.1%
% "General Interest Only"	(3.3%)	(2.0%)	(2.3%)	(0.8%)	(3.6%)	(1.1%)	(0.0%)	(3.7%)	(1.9%)	(3.9%)	(0.5%)
% "Major Elective"	(0.1%)	(0.0%)	(0.8%)	(0.6%)	(0.0%)	(0.0%)	(0.0%)	(1.7%)	(0.0%)	(0.1%)	(1.1%)
% "Major Requirement"	(2.4%)	(1.6%)	(2.5%)	(0.6%)	(2.0%)	(0.1%)	(0.0%)	(5.4%)	(0.5%)	(2.5%)	(2.3%)
% "General Ed Require"	(0.0%)	(0.1%)	(1.6%)	(0.2%)	(0.5%)	(0.2%)	(0.9%)	(0.0%)	(0.0%)	(0.8%)	(0.6%)
% "Minor Related Field"	(0.4%)	(0.2%)	(0.3%)	(0.0%)	(0.0%)	(0.0%)	(0.0%)	(1.0%)	(0.0%)	(0.3%)	(0.0%)
Year in School/Course Level[2]	0.0%	0.0%	0.0%	0.0%	(1.7%)	0.0%	1.5%	1.3%	1.3%	0.2%	1.1%
Course Level	(0.0%)	(0.0%)	(0.0%)	(0.2%)	(0.6%)	(0.1%)	(0.0%)	(0.2%)	(0.1%)	(0.0%)	(0.0%)
% Frosh-Soph in Class	(0.0%)	(0.0)	(0.0%)	(0.1%)	(0.0%)	(0.0%)	(1.2%)	(0.0%)	(0.0%)	(0.0%)	(0.1%)
% Jr-Sr in Class	(0.0%)	(0.0%)	(0.0%)	(0.0%)	(0.1%)	(0.0%)	(0.0%)	(0.0%)	(0.5%)	(0.1%)	(0.3%)
Avg. "Year in School"	(0.0%)	(0.0%)	(0.0%)	(0.3%)	(0.0%)	(0.0%)	(1.3%)	(0.1%)	(0.5%)	(0.0%)	(0.0%)
Overall GPA (prior)	1.0%	0.4%	1.0%	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	3.5%
% Division Majors	0.1%	0.5%	0.0%	0.0%	0.9%	1.1%	0.0%	0.1%	0.4%	0.0%	0.9%
Enrollment	0.0%	0.0%	0.0%	0.3%	0.0%	2.5%	1.1%	0.0%	0.2%	0.2%	0.0%
Teacher Rank	0.7%	0.0%	0.6%	1.0%	0.9%	1.0%	0.0%	1.8%	1.9%	0.8%	0.6%

[1]Uniqueness was defined as the adjusted change in R^2 due to the introduction of each Background/Demographic variable after all other variables are entered.

[2]For these two "sets" of variables, all variables were entered in one step and the change in adjusted R^2 for the entire set of variables is presented. Values in parentheses indicate the change in adjusted R^2 resulting form each variable separately, not considering other variables in the set. The contribution of one variable can be as high or higher than for the entire set since the R^2 is adjusted for the total number of variables included at each step.

TABLE FOUR

Canonical Variable Loadings for 511 Class Averages
(correlation between canonical variables and original variables)

Student Evaluation Factors	I	II	III	IV	V	
Overall Course	.64	.02	.28	.49	-.11	
Overall Instructor	.45	.02	.05	.38	-.02	
Learning	.78	-.12	.17	.16	-.23	
Enthusiasm	.40	-.13	.30	.07	-.14	
Organization	-.03	-.15	.03	.52	-.35	
Group Interaction	.64	-.35	-.34	.31	.17	
Individual Rapport	.24	-.17	-.15	.02	.02	
Breadth	.17	.54	-.27	.49	-.43	
Examinations	.22	-.19	.01	.76	-.01	
Assignments	.45	.30	.06	.45	-.34	
$\Sigma(ri^2)/n$	(.212)	(.062)	(.043)	(.179)	(.054)	
Background Variables						
Prior Interest	.79	-.10	.27	-.24	-.05	
Workload/Difficulty	.22	.49	.48	.35	.29	
Expected Grade	.55	-.34	-.27	.01	.04	
Course Level	.55	.09	-.35	-.05	.09	
% Major Elective	.53	.20	-.40	-.07	-.12	
% General Interest Only	.15	.06	.03	.32	-.64	
Overall GPA	.29	-.11	-.11	.23	-.43	
% Major Required	-.26	-.32	.11	-.08	.57	
% Majoring in Division	.32	-.05	-.40	.32	.39	
% Fosh Soph	-.38	.26	.37	.03	-.15	
% General Education	-.40	.21	.31	-.08	-.26	
Avg. Year in School	.45	-.27	-.29	-.29	.08	
Enrollment	-.31	.34	.48	-.17	-.25	
% Jr. Sr.	.19	-.15	-.33	.31	.22	
Teacher Renk	-.02	.76	-.41	-.20	.04	
% Minor/Related Field	.13	.18	-.04	.00	-.16	
$\Sigma(rj^2)/n$	(.156)	(.092)	(.014)	(.044)	(.089)	
Canonical R	.67	.61	.44	.42	.32	
Canonical R^2	.45	.37	.19	.18	.10	
Redundancy in Evaluations Explained by Background variables[1,2]	9.5%	2.3%	0.8%	3.2%	0.5%	Total Redundancy 16.3%
Redundancy in Background variables explained by Evaluations [1,2]	7.2%	3.4%	2.0%	0.8%	0.9%	14.1%

[1]Redundancy is an estimate of the total variance in one set of variables (evaluation scores) which is explained by another set of variables (background variables).

[2]Redundancy is defined as (R^2) $(\Sigma r^2/n)$ for each canonical variate. Total redundancy is the sum of these values for each of the canonical variates.

[3]The 16.3%, an estimate of the proportion of variance explained by the background variables, is positively biased. Although a correction for this bias has not been developed, the true population figure would be closer to the 11.8% estimated on the basis of the multiple regressions. (See Table Two)

TABLE FIVE

Preliminary Path Analysis: Evaluation Score Variance Explained By Four "Most Important" Background Variables [1]

Variables	Notation[2]	Over Crse.	Over Inst.	Lear.	Enth.	Org.	Group Int.	Ind. Rap.	Brea.	Exams	Assig.	Avg. Fig.
1. "Prior Subject Interest" unpartialed	$ry^2.1$	10.6%	4.1%	19.5%	5.2%	0.1%	8.3%	0.9%	0.1%	0.1%	3.9%	5.28%
2. "% Indicating General Interest" unpartialed	$ry^2.2$	2.7%	1.3%	2.3%	0.9%	2.4%	0.5%	0.0%	3.8%	1.0%	3.1%	1.80%
1. & 2. in Combination unpartialed	$R^2y.1,2$	12.7%	5.2%	21.2%	5.9%	2.6%	8.6%	0.9%	3.9%	1.0%	6.6%	6.86%
3. "Expected Grade" unpartialed	$R^2y.3$	4.6%	4.1%	8.3%	4.1%	0.0%	9.9%	3.0%	0.0%	3.3%	1.6%	3.89%
4. "Workload/Difficulty" unpartialed	$R^2y.4$	5.3%	1.8%	1.4%	0.4%	0.0%	0.0%	0.0%	2.4%	1.0%	5.37%	1.76%
3. & 4. in Combination unpartialed	$R^2y.3,4$	14.1%	8.5%	12.8%	5.7%	0.0%	11.9%	3.5%	2.5%	5.9%	9.5%	7.44%
3. with 1. & 2. partialed out	$R^2y.(3.1,2)$	2.2%	2.7%	3.9%	2.5%	0.0%	7.8%	2.5%	0.0%	3.2%	0.8%	2.56%
4. with 1. & 2. partialed out	$R^2y.(4.1,2)$	4.6%	1.6%	0.7%	0.2%	0.1%	0.2%	0.2%	3.5%	1.3%	5.5%	1.79%
3. & 4. in Comb. with 1. & 2. partialed out	$R^2y.(3,4.1,2)$	10.2%	6.6%	6.5%	3.6%	0.1%	8.1%	2.8%	3.8%	6.5%	8.6%	5.68%
Total Variance Explained	$R^2y.1,2,3,4$	22.9%	11.8%	27.7%	9.5%	2.7%	16.7%	3.7%	7.7%	7.5%	15.2%	12.54%

1 The model used assumes that (1) Prior Subject Interest & Reason for Taking Course precede (3) Expected Grade and (4) Workload/Difficulty

2 The notation used is that r^2 refers simple correlations $R^2_y.1,2$ refers to the total variance in each (y) evaluation score which is explained by Background variables (1) and (2); Ry. (3,4. 1,2) refers to the total variance in each evaluation score which is explained by Background variables (3) & (4) which is unique from that explained by variables (1) & (2).

3 None of the variance estimates are corrected for the number of variables in the regression equation. The Average of Total Variance Explained (12.54%) if corrected for 4 Background Variables would be 11.85% and 9.71% if corrected for the entire set of 16 variables.

APPENDIX II

RELIABILITY

Evaluation Items	ANOVA Reliability Estimates For Class[1] Averages Based Upon Different Numbers of Responses						Coefficient Alpha[2] Reliability Estimates of Factor Scores
	5	10	15	25	50	100	
I. LEARNING/VALUE							.95
Increased Interest as Course Consequence	.52	.69	.77	.83	.91	.96	
Learned Something Valuable	.55	.71	.78	.86	.92	.96	
Learned & Understood Subject Matter	.50	.67	.78	.85	.92	.95	
OVERALL COURSE RATING	.62	.76	.83	.89	.94	.96	
Intellectually Challenging/Stimulating	.64	.78	.84	.90	.95	.97	
II. ENTHUSIASM							.97
Dynamic & Energetic	.70	.83	.88	.92	.96	.98	
Enhanced with Humor	.69	.81	.87	.92	.96	.98	
Held Your Interest	.67	.80	.86	.91	.96	.97	
Enthusiastic About Teaching	.66	.79	.85	.91	.95	.97	
OVERALL INSTRUCTOR RATING	.66	.80	.85	.91	.95	.97	
III. ORGANIZATION							.93
Materials Prepared & Explained	.58	.74	.81	.88	.93	.97	
Instructor Explanations Clear	.60	.75	.82	.88	.94	.97	
Lectures Facilitated Note Taking	.60	.75	.82	.88	.94	.97	
Objectives Stated and Pursued	.51	.68	.76	.84	.91	.97	
IV. GROUP INTERACTION							.98
Students Shared Ideas/Knowledge	.64	.78	.84	.90	.95	.97	
Encouraged to Participate	.65	.79	.85	.90	.95	.97	
Encouraged to Express Own Ideas	.61	.76	.82	.89	.94	.97	
Encouraged to Question & Given Answers	.60	.75	.82	.88	.94	.97	
V. INDIVIDUAL RAPPORT							.95
Welcomed Seeking Help/Advice	.57	.72	.80	.87	.93	.96	
Interested in Individual Students	.57	.73	.80	.87	.93	.96	
Accessible to Students	.52	.69	.77	.85	.92	.96	
Friendly Toward Students	.57	.73	.80	.87	.93	.96	
VI. BREADTH OF COVERAGE							.93
Presented Background of Concepts	.55	.71	.78	.86	.92	.96	
Contrasted Implications	.52	.69	.77	.85	.92	.96	
Presented Different Points of View	.50	.67	.75	.83	.91	.95	
Discussed Current Developments	.56	.71	.79	.86	.94	.97	
VII. EXAMINATIONS							.94
Evaluation Methods Fair/Appropriate	.58	.74	.81	.88	.93	.97	
Tested Actual Content	.58	.74	.81	.88	.93	.97	
Exam Feedback Valuable	.59	.74	.81	.88	.94	.97	
VIII. ASSIGNMENTS							.90
Readings/Text Valuable	.63	.77	.84	.90	.94	.97	
Contributed to Understanding	.50	.67	.75	.83	.91	.95	
IX. WORKLOAD/DIFFICULTY							.88
Workload (Light-Heavy)	.60	.75	.82	.88	.94	.97	
Difficulty (Easy-Hard)	.52	.69	.77	.85	.92	.97	
Hours Out of Class	.55	.71	.78	.86	.92	.96	
Pace (Too Slow-Too Fast)	.36	.52	.62	.73	.85	.92	
MEDIAN RELIABILITY	.58	.74	.81	.88	.93	.97	.94

1--Anova Reliability estimates were obtained by taking 10 responses from each of 387 courses in which at least 15 students responded. A one-way Anova was performed in which the courses served as levels. The reliability estimate for 10 responses was computed by subtracting the reciprocal of the F-Ratio from 1.0. The other estimates were generated with the Spearman-Brown equation. This procedure is described in Winer (1971), Marsh (1976) and Centra (1973).

2--Coefficient Alphas were computed with Method 2 described by Nie, et. al. (1977).

Two types of reliability are presented above. The Anova reliability estimates measure the relative consistency within each class relative to the differences between different classes. The principle source of error measured by this technique is the diversity of student opinion within the courses. It should be noted that this is a more stringent criteria than would be measured by assessing the reliability of individual responses. Using the Spearman-Brown equation, the median reliability for a sample size of one would be $r = .22$. However, using a test-retest procedure over a three year interval, Overall and Marsh (1978) found that reliabilities of the responses of individual students were generally over .50.

The coefficient alpha reliability is based upon the degree of intercorrelation among the items defining each factor. This value will also vary with the number of responses. The average number of responses in the 511 courses used in this analysis was 26.7. (Avg. Enrollment was 34.56, Avg. Response Rate was 77%). The median reliability of the factor scores is substantially higher than the median reliability of individual items based upon a comparable number of responses. This is due, at least in part, to the greater reliability of an average.

APPENDIX III

Factor Analysis of Student Evaluation Instrument (N=511 Class Average Responses)

Factor Pattern Loadings

Evaluation Items (paraphrased)	Mean	Standard Deviation	I	II	III	IV	V	VI	VII	VIII	IX
I. LEARNING/VALUE											
Increased Interest as Course Consequence	3.91	0.56	69	14	-04	06	04	09	06	17	00
Learned Something Valuable	4.15	0.48	59	06	11	12	00	04	11	15	18
Learned & Understood Subject Matter	4.01	0.41	53	11	17	09	06	-09	10	12	-28
OVERALL COURSE RATING	3.83	0.65	44	23	12	07	05	07	20	17	10
Intellectually Challenging/Stimulating	3.98	0.54	43	17	03	08	-01	18	19	13	31
II. ENTHUSIASM											
Dynamic & Energetic	3.90	0.65	06	67	15	07	04	09	09	11	07
Enhanced with humor	3.85	0.65	01	67	16	00	08	06	07	10	00
Held your Interest	3.66	0.67	14	65	26	06	02	03	03	10	01
Enthusiatic about Teaching	4.18	0.57	10	48	19	07	13	14	13	09	06
OVERALL INSTRUCTOR RATING	3.97	0.65	14	43	25	10	17	11	12	08	05
III. ORGANIZATION											
Materials Prepared & Explained	3.90	0.56	12	-06	70	03	07	14	14	10	04
Instructor Explanations Clear	3.90	0.56	18	12	57	14	03	08	12	08	-07
Lectures Facilitated Note Taking	3.77	0.62	08	-02	51	-19	06	27	08	11	-03
Objectives stated & pursued	3.94	0.53	20	-10	49	06	08	12	24	11	04
IV. GROUP INTERACTION											
Students shared Ideas/knowledge	4.07	.59	08	10	-01	81	07	04	07	08	00
Encouraged to Participate	4.05	.60	11	12	03	80	07	00	00	07	00
Encouraged to Express Own Ideas	4.09	.55	06	12	04	73	16	07	11	04	00
Encouraged to Question & Given Answers	4.08	.55	09	13	17	62	16	04	12	07	-02
V. INDIVIDUAL RAPPORT											
Welcomed Seeking Help/Advice	4.13	.54	08	10	05	06	82	-02	10	03	-01
Interested in Individual Students	4.02	.57	06	18	06	17	69	-06	14	03	00
Accessible to Students	3.91	.56	-02	08	03	01	65	24	11	11	07
Friendly Towards Students	4.28	.49	00	-25	12	18	61	-08	10	06	-09
VI. BREADTH OF COVERAGE											
Presented Background of Concepts	3.97	.48	12	05	12	02	05	68	07	12	-03
Contrasted Implications	3.94	.49	06	10	03	07	08	67	01	20	04
Presented Different Points of View	4.03	.44	03	08	17	12	03	60	12	12	-03
Discussed Current Developments	4.14	.49	19	12	15	16	02	32	17	12	-06
VII. EXAMINATIONS											
Eval Methods Fair/Appropriate	3.80	.56	03	04	02	05	16	05	72	15	-06
Tested Actual Content	3.88	.55	09	02	10	02	06	09	67	14	-04
Exam Feedback Valuable	3.67	.59	03	05	09	10	16	-02	66	07	09
VIII. ASSIGNMENTS											
Readings/Text Valuable	3.72	.59	-02	-05	02	00	04	11	-01	91	02
Contributed to Understanding	3.88	.54	09	01	06	10	02	01	16	70	03
IX. WORKLOAD/DIFFICULTY											
Workload (Light-Heavy)	3.37	.61	10	02	02	08	00	00	00	08	89
Difficulty (Easy-Hard)	3.45	.52	-02	02	00	-01	-01	11	07	08	85
Hours Out of Class	2.61	.61	13	03	-10	01	10	07	-09	12	76
Pace (Too Slow-Too Fast)	3.09	.39	-09	01	11	-10	-05	12	14	04	60

1--Factor Analysis was Oblique (correlated) with the Delta Factor=2.0 (Nie, et. al, 1976)
2--First nine eigenvalues were 19.8, 3.3, 2.3, 1.5, 1.2, 1.0, .76, .60, .50
3--Correlations between Factors ranged from r=-.01 to r=.49 (Median r=.27)
4--All items except Workload/Difficulty were answered along 5-point response scale (1-Very Poor, 3-Moderate 5-Very Good). Workload/Difficult items varied on 5-point response scale with end-points above, except for Hours (1 to 2, 2-2 to 5, 3-5 to 7, 4-8 to 12.5-Over 12).

This is a sample of the SUMMARY REPORT that is sent to individual faculty and/or Department Chairpersons

INSTRUCTOR: DR. DOE J. COURSE: XXXXXXXX 999 CLASS SCHEDULE NUMBER: 99999 SEMESTER: FALL 76 PAGE 1 OF 2

NUMBER OF STUDENTS COMPLETING EVALUATIONS: 26 PERCENTAGE OF ENROLLED STUDENTS COMPLETING EVALUATIONS: 76%

STUDENT AND COURSE CHARACTERISTICS:

FOR EACH QUESTION THE PERCENTAGE OF STUDENTS MAKING EACH RESPONSE AND THE MEAN AVERAGE RESPONSE (IF APPROPRIATE) IS PRESENTED (THESE STATISTICS ARE BASED UPON THE NUMBER OF STUDENTS ACTUALLY RESPONDING TO THE ITEM). IN ADDITION, THE PERCENTAGE OF STUDENTS WHO COMPLETED THE EVALUATION FORM BUT DID NOT RESPOND TO A PARTICULAR QUESTION IS INDICATED BY THE "NO RESPONSE" PERCENTAGE.

PRIOR INTEREST	PCT
1-VERY LOW	0
2-	0
3-MEDIUM	40
4-	16
5-VERY HIGH	44
NO RESPONSE	3

OVERALL G.P.A.	PCT
1-BELOW 2.5	0
2-2.5 - 3.0	21
3-3.0 - 3.4	50
4-3.4 - 3.7	17
5-ABOVE 3.7	13
NO RESPONSE	7

EXPECTED GRADE	PCT
0-F	0
1-D	0
2-C	17
3-B	48
4-A	35
NO RESPONSE	11
MEAN: 3.2	

REASON IN CLASS	PCT
1-MAJ REQRD	40
2-MAJ ELECT	28
3-GEN ED REQ	4
4-MIN/RELTD	8
5-GEN INTRST	20
NO RESPONSE	3

YEAR IN SCHOOL	PCT
1-FRESHMAN	4
2-SOPHOMORE	12
3-JUNIOR	16
4-SENIOR	48
5-GRADUATE	20
NO RESPONSE	3

MAJOR DEPRTMNT	PCT
1-SOC SCI	65
2-NAT SCI	4
3-HUMNTIES	4
4-BUSINESS	0
6-ENGINEER	4
7-PERF ART	4
8-PUB AFFR	8
9-OTHER	0
0-UNDEC	12

COURSE DIFFICULTY	PCT
1-VERY EASY	4
2-	0
3-MEDIUM	77
4-	15
5-VERY HARD	4
NO RESPONSE	0
MEAN: 3.2	

COURSE WORKLOAD	PCT
1-VERY LIGHT	0
2-	0
3-MEDIUM	73
4-	15
5-VERY HEAVY	12
NO RESPONSE	0
MEAN: 3.4	

COURSE PACE	PCT
1-TOO SLOW	0
2-	0
3-RIGHT	85
4-	12
5-TOO FAST	4
NO RESPONSE	0
MEAN: 3.2	

OUTSIDE HRS/WK	PCT
1-0 TO 2	8
2-2 TO 5	35
3-5 TO 7	46
4-8 TO 12	12
5-OVER 12	0
NO RESPONSE	0
MEAN: 2.6	

SUMMARY EVALUATION SCORES

THESE SUMMARY SCORES (EXCEPT THE TWO OVERALL RATING ITEMS) ARE WEIGHTED AVERAGES OF SEPARATE EVALUATION ITEMS AND HAVE A MEAN AVERAGE (ACROSS ALL USC CLASSES) OF 50. THE STANDARD ERROR (SE) IS A MEASURE OF THE RELIABILITY OF EACH OF THE TWO OVERALL SUMMARY ITEMS. IT IS SMALLER (MORE RELIABLE) WHEN LARGER NUMBERS OF STUDENTS ARE RESPONDING AND WHERE THERE IS A GREATER AGREEMENT AMONG THE STUDENTS COMPLETING THE EVALUATIONS. DIFFERENCES OF LESS THAN ONE STANDARD ERROR ARE TOO SMALL TO BE RELIABLY INTERPRETED. IN GENERAL, EVALUATIONS BASED UPON LESS THAN 10 STUDENTS' RESPONSES OR EVALUATIONS BASED UPON LESS THAN 50% OF THE CLASS SHOULD BE INTERPRETED CAUTIOUSLY. THE PERCENTILE RANKS (WHICH VARY BETWEEN 0 & 100) AND THE GRAPHS SHOW HOW YOUR EVALUATIONS COMPARE WITH OTHER COURSES IN YOUR COMPARISON GROUP (HIGHER PERCENTILE RANKS AND MORE STARS INDICATE BETTER EVALUATIONS). YOUR COMPARISON GROUP IS: UNDERGRADUATE COURSES NOT TAUGHT BY TEACHING ASSISTANTS.

		MEAN	SE +/-	RANK RELATIVE TO YOUR COMPARISON GROUP (SEE ABOVE) XTIL RANK	GRAPH 0 1 2 3 4 5 6 7 8 9
LEARNING	VALUABLE LEARNING EXPERIENCE, WAS INTELLECTUALLY STIMULATING/CHALLENGING	63.2		93	**********
ENTHUSIASM	INSTR DISPLAYED ENTHUSIASM, ENERGY, HUMOR & ABILITY TO HOLD INTEREST	46.3		28	******
ORGANIZATION	ORGANIZATION/CLARITY OF EXPLANATIONS, COURSE MATERIALS, OBJECTIVES, LECTURES	43.7		21	*****
GROUP INTERACT	STUDENTS ENCOURAGED TO DISCUSS, PARTICIPATE, SHARE IDEAS & ASK QUESTIONS	38.7		16	****
INDV. RAPPORT	INSTRUCTOR ACCESIBLE, FRIENDLY, AND INTERESTED IN STUDENTS	39.7		13	***
BREADTH	PRESENTATION OF BROAD BACKGRD, CONCEPTS & ALTERNATIVE APPROACHES/THEORIES	50.8		46	**********
EXAMINATIONS	STUDENT PECEPTIONS OF VALUE & FAIRNESS OF EXAMS/GRADED MATERIALS	50.8		49	**********
ASSIGNMENTS	VALUE OF ASSIGNMENTS IN ADDING APPRECIATION/UNDERSTANDING TO COURSE	53.6		62	**********
OVERALL COURSE	HOW DOES THIS COURSE COMPARE WITH OTHERS AT U.S.C.? (QUESTION 30)	3.8	0.2	46	**********
OVERALL INSTR.	HOW DOES THIS INSTRUCTOR COMPARE WITH OTHERS AT U.S.C.? (QUESTION 31)	3.8	0.2	37	********

PREPARED BY THE OFFICE OF INSTITUTIONAL STUDIES, USC; TELEPHONE: (213)746-6503.
INSTRUCTOR: DR. DOE J. COURSE: XXXXXXXX 999 CLASS SCHEDULE NUMBER: 99999 SEMESTER: FALL 76 PAGE 2 OF 2

NUMBER OF STUDENTS COMPLETING EVALUATIONS: 26 PERCENTAGE OF ENROLLED STUDENTS COMPLETING EVALUATIONS: 76%

EVALUATION ITEMS (SOME QUESTIONS HAVE BEEN ABBREVIATED):

FOR EACH QUESTION, THE PERCENTAGE OF STUDENTS MAKING EACH RESPONSE, THE MEAN AVERAGE RESPONSE, AND THE STANDARD ERROR (SE) OF THE RESPONSES ARE PRESENTED (THESE STATISTICS ARE BASED UPON THE ACTUAL NUMBER OF STUDENTS RESPONDING TO THE QUESTION). IN ADDITION, THE PERCENTAGE OF STUDENTS WHO COMPLETED THE EVALUATION FORM BUT DID NOT RESPOND TO A PARTICULAR QUESTION IS INDICATED IN THE "NO RESP" COLUMN. DIFFERENCES IN MEAN AVERAGES THAT ARE LESS THAN ONE STANDARD ERROR (SEE PAGE ONE FOR A DESCRIPTION) ARE TOO SMALL TO BE RELIABLY INTERPRETED. IN GENERAL, EVALUATIONS BASED UPON LESS THAN 10 STUDENTS' RESPONSES, EVALUATIONS BASED UPON LESS THAN 50% OF THE CLASS, AND EVALUATION ITEMS WHICH WERE FREQUENTLY LEFT BLANK SHOULD BE INTERPRETED CAUTIOUSLY. THE PERCENTILE RANKS (WHICH VARY BETWEEN 0 & 100) AND THE GRAPHS SHOW HOW YOUR EVALUATIONS COMPARE WITH OTHER COURSES IN YOUR COMPARISON GROUP (HIGHER PERCENTILE RANKS AND MORE STARS INDICATE MORE FAVORABLE EVALUATIONS). YOUR COMPARISON GROUP IS: UNDERGRADUATE COURSES NOT TAUGHT BY TEACHING ASSISTANTS.

	PERCENTAGE RESPONDING								RANK RELATIVE TO YOUR COMPARISON GROUP (SEE ABOVE)	
	VERY POOR	POOR	MED-IUM	GOOD	VERY GOOD	NO RESP	MEAN	SE +/-	%TIL RANK	GRAPH 0 1 2 3 4 5 6 7 8 9
LEARNING										
1. CRSE WAS INTELLECTUALLY CHALLENGING AND STIMULATING	0	0	15	54	31	0	4.2	0.13	64	**********
2. LEARNED SOMETHING CONSIDERED TO BE VALUABLE	0	0	4	46	50	0	4.5	0.11	73	**********
3. INCREASED INTEREST IN SUBJECT AS CONSEQUENCE OF CRSE	0	0	12	31	58	0	4.5	0.14	84	**********
4. LEARNED AND UNDERSTOOD THE SUBJECT MATERIALS	0	4	12	64	20	3	4.0	0.14	52	**********
ENTHUSIASM										
5. INSTR WAS ENTHUSIASTIC ABOUT TEACHING THE CRSE	0	8	23	38	31	0	3.9	0.18	29	******
6. INSTR WAS DYNAMIC & ENERGETIC IN CONDUCTING CRSE	0	16	32	36	16	3	3.5	0.19	26	******
7. INSTR ENHANCED PRESENTATION WITH HUMOR	4	4	15	42	35	0	4.0	0.20	60	**********
8. INSTR STYLE OF PRESENTATION HELD INTEREST	4	15	23	42	15	0	3.5	0.21	41	*********
ORGANIZATION										
9. INSTR'S EXPLANATIONS WERE CLEAR	0	13	29	46	13	7	3.6	0.18	26	******
10. CRSE MATERIALS WERE WELL PREPARED & EXPLAINED	4	8	23	54	12	0	3.6	0.18	29	******
11. PROPOSED OBJECTIVES AGREED WITH THOSE ACTUALLY TAUGHT	0	19	31	31	19	0	3.5	0.20	20	*****
12. LECTURES FACILITATED TAKING NOTES	0	4	31	54	12	0	3.7	0.14	45	**********
GROUP INTERACTION										
13. STUDENTS ENCOURAGED TO PARTICIPATE IN CLASS DISCUSSIONS	8	0	27	46	19	0	3.7	0.21	29	******
14. STUDENTS INVITED TO SHARE IDEAS & KNOWLEDGE	4	8	35	42	12	0	3.5	0.19	23	*****
15. STUDENTS ENCOURAGED TO ASK QUESTIONS AND GIVE ANSWERS	4	8	24	56	8	3	3.6	0.18	17	****
16. STUDENTS ENCOURAGED TO EXPRESS OWN IDEAS	8	8	31	46	8	0	3.4	0.20	16	****
INDIVIDUAL RAPPORT										
17. INSTR WAS FRIENDLY TOWARDS INDIVIDUAL STUDENTS	4	8	15	46	27	0	3.8	0.21	17	****
18. INSTR WELCOMED STUDENTS TO SEEK HELP/ADVICE	0	15	27	38	19	0	3.6	0.19	17	****
19. INSTR HAD GENUINE INTEREST IN INDIVIDUAL STUDENTS	4	4	40	40	12	3	3.5	0.18	21	*****
20. INSTR WAS ACCESSIBLE DURING OFFICE HRS/AFTER CLASS	4	8	48	40	0	3	3.2	0.16	10	***
BREADTH										
21. INSTR CONTRASTED IMPLICATIONS OF VARIOUS THEORIES	4	8	19	42	27	0	3.8	0.21	38	********
22. INSTR PRESENTED BACKGROUND OF IDEAS/CONCEPTS	0	0	20	44	36	3	4.2	0.15	64	**********
23. INSTR PRESENTED POINTS OF VIEW OTHER THAN OWN	0	4	19	38	38	0	4.1	0.17	56	**********
24. INSTR DISCUSSED CURRENT DEVELOPMENTS IN FIELD	0	8	4	46	42	0	4.2	0.17	57	**********
EXAMINATIONS										
25. FEEDBACK ON EXAMS/GRADED MATERIALS WAS VALUABLE	4	4	35	27	31	0	3.8	0.21	56	**********
26. METHOD OF EVALUATION WAS FAIR & APPROPRIATE	0	0	31	38	31	0	4.0	0.16	66	**********
27. GRADED MATERIALS TESTED CRSE CONTENT AS EMPHASIZED	0	16	24	32	28	3	3.7	0.21	36	********
ASSIGNMENTS										
28. REQUIRED READINGS/TEXTS WERE VALUABLE	0	8	27	31	35	0	3.9	0.19	57	**********
29. ASSIGNMENTS CONTRIBUTED TO APPRECIATION/UNDERSTANDING	0	4	15	50	31	0	4.1	0.16	62	**********

USC EVALUATION SERVICES

PUT INSTRUCTOR'S NAME & COMMENTS ON BACK—USE PENCIL ONLY, NOT PEN—COMPLETELY ERASE ANY RESPONSES YOU CHANGE

AS A DESCRIPTION OF THIS COURSE/INSTRUCTOR, THIS STATEMENT IS:
(SELECT THE BEST RESPONSE FOR EACH OF THE FOLLOWING STATEMENTS. LEAVING A RESPONSE BLANK ONLY IF IT IS CLEARLY NOT RELEVANT)

		VERY POOR	POOR	MEDIOCRE	GOOD	VERY GOOD
1	**LEARNING:** YOU FOUND THE COURSE INTELLECTUALLY CHALLENGING AND STIMULATING	1	2	3	4	5
2	YOU HAVE LEARNED SOMETHING WHICH YOU CONSIDER VALUABLE	1	2	3	4	5
3	YOUR INTEREST IN THE SUBJECT HAS INCREASED AS A CONSEQUENCE OF THIS COURSE	1	2	3	4	5
4	YOU HAVE LEARNED AND UNDERSTOOD THE SUBJECT MATERIALS IN THIS COURSE	1	2	3	4	5
5	**ENTHUSIASM:** INSTRUCTOR WAS ENTHUSIASTIC ABOUT TEACHING THE COURSE	1	2	3	4	5
6	INSTRUCTOR WAS DYNAMIC AND ENERGETIC IN CONDUCTING THE COURSE	1	2	3	4	5
7	INSTRUCTOR ENHANCED PRESENTATIONS WITH THE USE OF HUMOR	1	2	3	4	5
8	INSTRUCTOR'S STYLE OF PRESENTATION HELD YOUR INTEREST DURING CLASS	1	2	3	4	5
9	**ORGANIZATION:** INSTRUCTOR'S EXPLANATIONS WERE CLEAR	1	2	3	4	5
10	COURSE MATERIALS WERE WELL PREPARED AND CAREFULLY EXPLAINED	1	2	3	4	5
11	PROPOSED OBJECTIVES AGREED WITH THOSE ACTUALLY TAUGHT SO YOU KNEW WHERE COURSE WAS GOING	1	2	3	4	5
12	INSTRUCTOR GAVE LECTURES THAT FACILITATED TAKING NOTES	1	2	3	4	5
13	**GROUP INTERACTION:** STUDENTS WERE ENCOURAGED TO PARTICIPATE IN CLASS DISCUSSIONS	1	2	3	4	5
14	STUDENTS WERE INVITED TO SHARE THEIR IDEAS AND KNOWLEDGE	1	2	3	4	5
15	STUDENTS WERE ENCOURAGED TO ASK QUESTIONS & WERE GIVEN MEANINGFUL ANSWERS	1	2	3	4	5
16	STUDENTS WERE ENCOURAGED TO EXPRESS THEIR OWN IDEAS AND/OR QUESTION THE INSTRUCTOR	1	2	3	4	5
17	**INDIVIDUAL RAPPORT:** INSTRUCTOR WAS FRIENDLY TOWARDS INDIVIDUAL STUDENTS	1	2	3	4	5
18	INSTRUCTOR MADE STUDENTS FEEL WELCOME IN SEEKING HELP/ADVICE IN OR OUTSIDE OF CLASS	1	2	3	4	5
19	INSTRUCTOR HAD A GENUINE INTEREST IN INDIVIDUAL STUDENTS	1	2	3	4	5
20	INSTRUCTOR WAS ADEQUATELY ACCESSIBLE TO STUDENTS DURING OFFICE HOURS OR AFTER CLASS	1	2	3	4	5
21	**BREADTH:** INSTRUCTOR CONTRASTED THE IMPLICATIONS OF VARIOUS THEORIES	1	2	3	4	5
22	INSTRUCTOR PRESENTED THE BACKGROUND OR ORIGIN OF IDEAS/CONCEPTS DEVELOPED IN CLASS	1	2	3	4	5
23	INSTRUCTOR PRESENTED POINTS OF VIEW OTHER THAN HIS/HER OWN WHEN APPROPRIATE	1	2	3	4	5
24	INSTRUCTOR ADEQUATELY DISCUSSED CURRENT DEVELOPMENTS IN THE FIELD	1	2	3	4	5
25	**EXAMINATIONS:** FEEDBACK ON EXAMINATIONS/GRADED MATERIALS WAS VALUABLE	1	2	3	4	5
26	METHODS OF EVALUATING STUDENT WORK WERE FAIR AND APPROPRIATE	1	2	3	4	5
27	EXAMINATIONS/GRADED MATERIALS TESTED COURSE CONTENT AS EMPHASIZED BY THE INSTRUCTOR	1	2	3	4	5
28	ASSIGNMENTS: REQUIRED READINGS/TEXTS WERE VALUABLE	1	2	3	4	5
29	READINGS, HOMEWORK, ETC. CONTRIBUTED TO APPRECIATION AND UNDERSTANDING OF SUBJECT	1	2	3	4	5
30	OVERALL: HOW DOES THIS COURSE COMPARE WITH OTHER COURSES YOU HAVE HAD AT USC?	1	2	3	4	5
31	HOW DOES THIS INSTRUCTOR COMPARE WITH OTHER INSTRUCTORS YOU HAVE HAD AT USC?	1	2	3	4	5

STUDENT AND COURSE CHARACTERISTICS (LEAVE BLANK IF NO RESPONSE APPLIES)

32	COURSE DIFFICULTY, RELATIVE TO OTHER COURSES, WAS (1-VERY EASY ... 3-MEDIUM ... 5-VERY HARD)	1	2	3	4	5
33	COURSE WORKLOAD (RELATIVE TO OTHER COURSES) WAS (1-VERY LIGHT ... 3-MEDIUM ... 5-VERY HEAVY)	1	2	3	4	5
34	COURSE PACE WAS (1-TOO SLOW ... 3-ABOUT RIGHT ... 5-TOO FAST)	1	2	3	4	5
35	HOURS/WEEK REQUIRED OUTSIDE OF CLASS: 1: 0 TO 3 ... 2: 3 TO 5 ... 3: 5 TO 7 ... 4: 9 TO 12 ... 5: OVER 12	1	2	3	4	5
36	LEVEL OF INTEREST IN THE SUBJECT PRIOR TO THIS COURSE (1-VERY LOW ... 3-MEDIUM ... 5-VERY HIGH)	1	2	3	4	5
37	OVERALL GPA AT USC (1: BELOW 2.5 ... 2: 2.5 TO 2.9 ... 3: 3.0 TO 3.4 ... 4: 3.4 TO 3.7 ... 5: ABOVE 3.7 LEAVE BLANK IF NOT YET ESTABLISHED AT USC)	1	2	3	4	5
38	EXPECTED GRADE IN THE COURSE (1-F, 2-D, 3-C, 4-B, 5-A)	F	D	C	B	A
39	REASON FOR TAKING THE COURSE: 1-MAJOR REQUIRE ... 2-MAJOR ELECTIVE ... 3-GENERAL ED REQUIRE 4-MINOR/RELATED FIELD ... 5-GENERAL INTEREST ONLY ... SELECT THE ONE WHICH IS BEST	1	2	3	4	5
40	YEAR IN SCHOOL: 1: FRESH ... 2: SOPH ... 3: JR ... 4: SR ... 5: GRAD	1	2	3	4	5
41	MAJOR DEPARTMENT: 1: SOC SCI/COMM ... 2: NAT SCI/MATH ... 3: HUMANITIES ... 4: BUSINESS ... 5: EDUCATION 6: ENGINEERING ... 7: PERF ARTS ... 8: PUB AFFAIRS ... 9: OTHER ... 10: UNDECLARED/UNDECIDED	1	2	3	4	5
		6	7	8	9	10

SUPPLEMENTAL QUESTIONS (USE RESPONSES BELOW FOR INSTRUCTOR'S QUESTIONS)

42	1	2	3	4	5	47	1	2	3	4	5	52	1	2	3	4	5	57	1	2	3	4	5
43	1	2	3	4	5	48	1	2	3	4	5	53	1	2	3	4	5	58	1	2	3	4	5
44	1	2	3	4	5	49	1	2	3	4	5	54	1	2	3	4	5	59	1	2	3	4	5
45	1	2	3	4	5	50	1	2	3	4	5	55	1	2	3	4	5	60	1	2	3	4	5
46	1	2	3	4	5	51	1	2	3	4	5	56	1	2	3	4	5	61	1	2	3	4	5

© 1976 HERBERT W. MARSH OFFICE OF INSTITUTIONAL STUDIES, UNIVERSITY OF SOUTHERN CALIFORNIA, L.A. 90007 IBM #58495

INSTRUCTOR'S NAME	DEPARTMENT NAME	COURSE NUMBER

INSTRUCTIONS

This evaluation form is intended to measure your reactions to this instructor and course. Results will be reported to the Department Chairmen to be used as part of the overall evaluation of the instructor. These evaluations will have budgetary and promotional implications so please take it very seriously. When you have finished a designated student will pick up the evaluations and take them to the Department Chairperson. Your responses will remain anonymous and summaries will not be given to the instructor until after the final grades have been assigned.

****Put Instructor's Name, Department Name and Course Number at top
(i.e., Smith, Psychology, 200)
****Use a number 2 pencil, do not use ink, ball point, magic marker, etc.
****Blacken only one response for each question and erase any changes completely

OPEN-ENDED COMMENTS

PLEASE INDICATE THE IMPORTANT CHARACTERISTICS OF THIS INSTRUCTOR/COURSE WHICH HAVE BEEN MOST VALUABLE TO YOUR LEARNING EXPERIENCE.

1.

2.

3.

PLEASE INDICATE CHARACTERISTICS OF THIS INSTRUCTOR/COURSE WHICH YOU FEEL ARE MOST IMPORTANT FOR HIM/HER TO WORK ON IMPROVING (PARTICULARLY ASPECTS NOT COVERED BY RATING ITEMS).

1.

2.

3.

PLEASE USE THE ADDITIONAL SPACE TO CLARIFY ANY OF YOUR RESPONSES OR TO MAKE OTHER COMMENTS.

IBM-H59496